(51) International Patent Classification⁷: G06F 17/18

(21) International Application Number:
PCT/AU2003/000923

(22) International Filing Date: 18 July 2003 (18.07.2003)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant (for all designated States except US): COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION [AU/AU]; Limestone Avenue, CAMPBELL, Australian Capital Territory 2612 (AU).

(72) Inventor; and
(75) Inventor/Applicant (for US only): STONE, Glenn [AU/AU]; 19/55-61 Old Northern Road, BAULKHAM HILLS, New South Wales 2153 (AU).
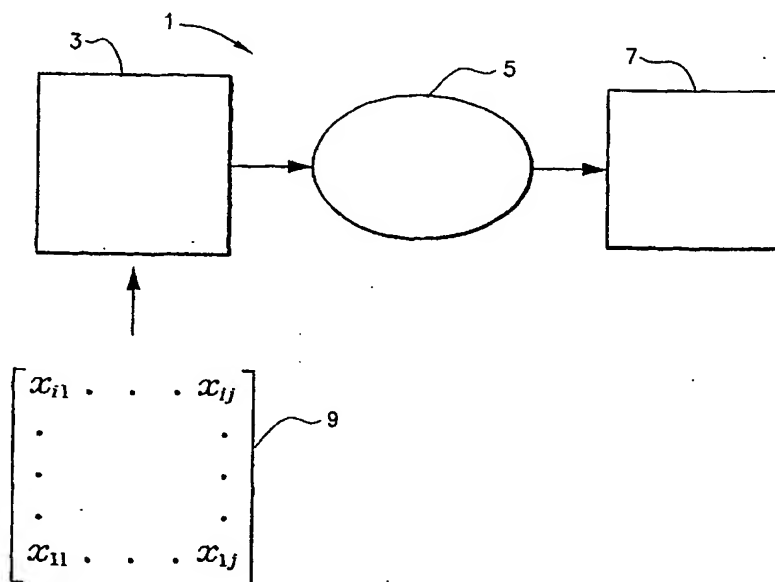
(74) Agent: GRIFFITH HACK; GPO Box 4164, SYDNEY, New South Wales 2001 (AU).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: A METHOD AND SYSTEM FOR SELECTING ONE OR MORE VARIABLES FOR USE WITH A STATISTICAL MODEL

$$\begin{bmatrix} x_{i1} & . & . & . & x_{ij} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ x_{11} & . & . & . & x_{1j} \end{bmatrix} \sim 9$$

(57) Abstract: A method of selecting one or more variables for use with a statistical model, the method comprising the steps of: creating a plurality of unique subsets of variables of multivariate data; determining the performance of a discriminant rule when used with each of the subsets, the discriminant rule being based on multivariate normal class densities each having substantially diagonal covariance matrices; and selecting the one or more variables from at least one of the subsets that result in a desired performance of the discriminant rule.